

UC San Diego

UC San Diego Previously Published Works

Title

Enteric dysbiosis and fecal calprotectin expression in premature infants.

Permalink

<https://escholarship.org/uc/item/5k7124c8>

Journal

Pediatric research, 85(3)

ISSN

0031-3998

Authors

Ho, Thao TB
Groer, Maureen W
Kane, Bradley
et al.

Publication Date

2019-02-01

DOI

10.1038/s41390-018-0254-y

Peer reviewed



Published in final edited form as:

Pediatr Res. 2019 February ; 85(3): 361–368. doi:10.1038/s41390-018-0254-y.

Enteric Dysbiosis and Fecal Calprotectin Expression in Premature Infants

Thao T. B. Ho¹, Maureen W. Groer^{1,2}, Bradley Kane², Alyson L. Yee^{3,4,5}, Benjamin A. Torres¹, Jack A. Gilbert^{4,5,6}, and Akhil Maheshwari^{7,*}

¹Department of Pediatrics, Morsani College of Medicine, University of South Florida, Tampa, Florida. ²College of Nursing, University of South Florida, Tampa, Florida. ³Interdisciplinary Scientist Training Program, University of Chicago, Chicago, Illinois. ⁴Microbiome Center, University of Chicago, Chicago, Illinois. ⁵Department of Surgery, University of Chicago, Chicago, Illinois. ⁶Argonne National Laboratory, Chicago, Illinois. ⁷Department of Pediatrics, Johns Hopkins University School of Medicine, Baltimore, Maryland.

Abstract

Background: Premature infants often develop enteric dysbiosis with a preponderance of Gammaproteobacteria, which has been related to adverse clinical outcomes. We investigated the relationship between increasing fecal Gammaproteobacteria and mucosal inflammation, measured by fecal calprotectin (FC).

Methods: Stool samples were collected from very-low-birth weight (VLBW) infants at 2, 3, and 4 weeks' postnatal age. Fecal microbiome was surveyed using PCR-amplification of the V4 region of 16S rRNA, and FC was measured by enzyme immunoassay.

Results: We enrolled 45 VLBW infants (gestation 27.9±2.2 weeks, birth weight 1126±208 g) and obtained stool samples at 9.9±3, 20.7±4.1, and 29.4±4.9 days. FC was positively correlated with the genus *Klebsiella* ($r = 0.207$, $p = 0.034$) and its dominant amplicon sequence variant ($r = 0.290$, $p = 0.003$) but not with the relative abundance of total Gammaproteobacteria. *Klebsiella* colonized the gut in two distinct patterns: some infants started with low *Klebsiella* abundance and gained these bacteria over time, whereas others began with very high *Klebsiella* abundance.

Users may view, print, copy, and download text and data-mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use: http://www.nature.com/authors/editorial_policies/license.html#terms

*Corresponding Author: Akhil Maheshwari, MD, 1800 Orleans St., Charlotte R. Bloomberg Children's Center Rm 8530, Baltimore, MD 21287, Phone: 708-910-8729; akhil@jhmi.edu.

Authors' contributions

TTBH collected and analyzed the data and wrote the manuscript. MWG supervised laboratory process and advised on content of the manuscript. BK performed laboratory analysis and advised on laboratory techniques. ALY analyzed 16S rRNA data. BAT advised on the content of the manuscript. JAG supervised the analysis of 16S rRNA data and advised on the manuscript content. AM supervised the data analysis and wrote the manuscript. All authors edited the manuscript and approved the final draft.

Competing interests

The authors declare that they have no competing interests to disclose.

Disclosure Statement: The authors declare that they have no competing interests to disclose.

Category of Study: This is a clinical study

Conclusion: In premature infants, FC correlated with relative abundance of a specific pathobiont, *Klebsiella*, and not with that of the class Gammaproteobacteria. These findings indicate a need to define dysbiosis at genera or higher levels of resolution.

BACKGROUND

In healthy, full-term neonates, the intestine is seeded by microbial inocula received from the birth canal, maternal skin, and milk, and the gut microbiome is dominated by Gram-positive Firmicutes such as *Staphylococcus*, *Propionibacterium*, *Bifidobacterium*, and *Lactobacillus* (1). Increasing pre-clinical and epidemiological information now acknowledges the positive influence of these bacteria on mucosal homeostasis, immunity, nutrient absorption, and energy regulation in the neonatal intestine (2,3). In contrast, in preterm infants, the intestinal microbiome may develop very differently (4). In these infants, the hospital microflora may contribute more to the gut microbiome, whereas the maternal influence may be curtailed due to shortened labor or a Cesarean birth, limited physical contact with the mother, and periods of enteral fasting. There may be additional selection pressures from broad-spectrum antibiotics, indwelling tubes and catheters, intestinal dysmotility, and immaturity of barriers such as gastric acid, secretory IgA, and mucus (1). Not surprisingly, many preterm infants develop an enteric dysbiosis with a preponderance of Gram-negative bacteria of the class Gammaproteobacteria and its constituent families *Enterobacteriaceae*, *Vibrionaceae*, and *Pseudomonadaceae* (5–8). Emerging epidemiological data link such dysbiosis with adverse outcomes such as necrotizing enterocolitis, cholestasis, and developmental delay (4,6,7,9), but the mechanism(s) by which Gammaproteobacteria present in the bowel lumen may evoke inflammatory responses are not well-understood. Intestinal epithelium and many resident immune cell lineages are normally tolerant of Gram-negative bacteria (10), and further study is needed to ascertain if dysbiosis secondary to a gammaproteobacterial bloom may induce mucosal inflammation when the relative abundance of Gammaproteobacteria exceeds a critical threshold, or is associated with an increase in specific pathobionts in this class of bacteria.

In this prospective observational study, we investigated the relationship between fecal Gammaproteobacteria and mucosal inflammation in premature infants. We used fecal calprotectin (FC), which is derived from neutrophils and inflammatory macrophages present in the mucosa, as the inflammatory marker (11). Preterm infants tend to have greater variability in FC expression than older children and adults, but several studies now show FC to be a useful marker of intestinal inflammation and necrotizing enterocolitis (NEC) when cut-off values between 200–300 µg/g of stool were used (11–14). We hypothesized that that FC increases in preterm infants as a function of the relative abundance of Gammaproteobacteria in stool. To test this hypothesis, we recorded demographic and clinical information from a cohort of very-low-birth weight (VLBW) infants, surveyed their fecal microbiome, and measured FC at serial time-points during the neonatal period.

METHODS:

Demographic and clinical information:

This study was performed at University of South Florida and the affiliated Tampa General Hospital (TGH) under appropriate oversight from the Institutional Review Boards. The neonatal intensive care unit (NICU) at TGH is an academic regional referral center with a single-patient room floor plan. All eligible VLBW infants were enrolled during the period May 2012 to December 2013; inclusion criteria included birth weight <1500 g, availability of a stool sample in the first 2 weeks after birth, and parental informed consent. Infants with major congenital anomalies or lethal chromosomal disorders were excluded. Three stool samples were obtained from each infant at 2, 3rd, and the 4th week after birth and stored at -80°C until the time of analysis. Biological samples were handled under uniform conditions per accepted guidelines (biospecimen reporting for improved study quality). (15)

Fecal DNA amplification:

Total DNA was extracted from 100–250 mg of stool (MoBio PowerFecal DNA kit, Qiagen, Carlsbad, CA) and V4 region of bacterial 16S rRNA gene was amplified by polymerase chain reaction with modified 515F and 806R primers (16). Amplicons were sequenced according to Illumina MiSeq protocol (Illumina, San Diego, CA) to generate about 15,000 of 250 base-pair paired-end reads per sample.

Metagenomic analysis:

Demultiplexed DNA sequences were first uploaded to the CLC Biomedical Workbench 3.5.3 (Qiagen) and analyzed for bacterial classifications to genus level. Bacterial Operational Taxonomic Units (OTUs) were mapped to the Greengenes v13.8 reference database at 97% similarity, and their relative abundances (%) were computed. Diversity within samples (alpha diversity) were measured and represented by the number of OTUs, phylodiversity, Chao 1, Simpson, and Shannon indices. To improve taxonomic resolution, we next used the Divisive Amplicon De-noising Algorithm 2 (DADA2) to characterize bacterial OTUs to amplicon sequence variants (ASVs). This 2-step analysis provided information on OTUs to allow comparison with existing studies of the preterm microbiome (5–7,17–19), and allowed validation of our results with the DADA2 pipeline that relies on a different, model-based approach to minimize amplicon errors.

Changes in the abundance of individual OTUs were investigated using Gneiss (balance tree) analysis (20). Due to the compositional nature of 16S rRNA gene sequence data, the relative (proportional) abundance of a taxon could be altered either due to a change in its own numbers or due to a shift in the relative abundance of other taxa. Balance tree analysis circumvents this problem by leveraging the correlation between co-occurring taxa to infer meaningful properties of sub-communities, rather than individual taxa.

Fecal Calprotectin (FC) measurements:

We used a commercially-available enzyme immunoassay (PhiCal, Genova Diagnostics, Ashville, NC). The linear range of measurement for this assay is 25–2500 µg/g stool.

Statistical Analysis:

Clinical and demographic data were analyzed using SPSS 25 (IBM, Armonk, NY). Scalar variables were compared by Friedman's test for repeated measures or the Mann-Whitney *U* test, and categorical variables by Fisher's exact test. Linear mixed-effects models were used to identify determinants of FC and *Klebsiella* abundance. Small-for-gestational age (SGA), ethnicity, vaginal birth, antenatal steroids, magnesium sulfate, chorioamnionitis, gender, multiple births, postnatal age, enteral feedings (binary), respiratory distress syndrome, patency of ductus arteriosus, sepsis, red blood cell transfusions (binary), and *Klebsiella* abundance were defined as fixed effects as all possible levels of the variable were represented in the cohort. Maternal body-mass index (BMI), birth weight, gestation, postmenstrual age at stool collection, duration of ruptured membranes, and days of antibiotic treatment were defined as random effects. Mixed-effects procedures were performed using the maximum likelihood method and an autoregressive covariance matrix (with heterogeneous variances) as the dependent variables were anticipated to diverge with time. Best-fitting models were identified by $-2 \log$ likelihood, Akaike's information criterion, and Schwarz's Bayesian criterion. Important independent variables were shortlisted using bootstrap bagging, where a bootstrap dataset was constructed by not sampling a third of all subjects and replacing these by an equal number of duplicated samples. The bootstrap sample was analyzed by logistic regression with entry criterion of $p < 0.2$. To avoid model overfitting, multivariable analyses were limited to biologically plausible associations, to main effects for baseline measures, and time-dependent covariates for longitudinal measures.

We also performed linear regression to identify predictors of FC at different time-points of stool collection. Independent variables were tested with entry at $p < 0.2$ and acceptance at $p < 0.05$, first using a onestep forced entry and then 'stepwise' in the sequence of appearance during perinatal period. In addition to the variables listed above, duration of iron therapy was also tested. To identify highly-correlated variables, multicollinearity diagnostics (tolerance values < 0.2 , variance inflation factors > 10) were reviewed. Independence of variables was confirmed by the Durbin-Watson statistic (acceptable range 1.5–2.5). Scatterplots of standardized residuals vs. standardized predicted values were evaluated for homoscedasticity and non-linearity. Normality of residuals was confirmed by evaluation of histograms. To determine *Klebsiella* abundance levels associated with elevated FC, we computed receiver-operating characteristics (ROCs) by plotting sensitivity vs. 1-specificity and identified 'cut-off' abundance values at the highest sum of sensitivity and specificity (Youden's *J* statistic). Statistical tests were two-tailed and considered significant at $p < 0.05$.

RESULTS

Demographic and clinical information:

We enrolled 45 eligible VLBW infants admitted to our NICU between May 2012 and December 2013. The clinical characteristics of this cohort have been described in part, elsewhere (21). These infants were born at a gestation (mean \pm standard deviation, SD) of 27.9 ± 2.2 weeks, with birth weight 1126 ± 208 g. Their clinical characteristics are summarized in table 1 of our previous publication (22).

Fecal calprotectin:

Stool samples for the 2, 3rd, and 4th week time-points were collected at postnatal age (mean \pm SD) 9.9 \pm 3, 20.7 \pm 4.1, and 29.4 \pm 4.9 days; and post-menstrual age 29.8 \pm 2.3, 31.2 \pm 1.9, and 32.6 \pm 1.9 weeks, respectively. FC levels remained stable over time [median (interquartile range/IQR): 161 (101–277), 187 (132–306), and 152 (102–311) μ g/g stool in 2, 3rd and 4th week, respectively; Friedman's test; difference not significant (*N.S.*)].

Fecal microbiome:

A total of 2,017,727 reads was obtained (mean \pm SD 15,285 \pm 7,139 sequences per sample). The alpha-diversity metrics (number of OTUs, phylodiversity, and Shannon, Chao1, and Simpson indices) increased with postnatal age (Figure 1). The relative abundance of Proteobacteria increased over time, comprising 46% (median; IQR = 0–90%) of all reads at 2 weeks, 83.5% (54.8–93.3%) in the 3rd, and 77% (57–88.3%) in the 4th week ($p < 0.001$). Gammaproteobacteria accounted for 42.5% (0–90%) of all reads in 2 weeks, 69.7% (29.9–86.9%) in the 3rd, and 75.5% (54.5–86%) in the 4th week stool samples ($p < 0.001$). *Klebsiella* were the dominant gammaproteobacterial genus: median (IQR) 44% (2–98%), 85% (0–98%), and 78.5% (11.3–97.5%) in the 2, 3rd, and 4th week stool samples, respectively as summarized in supplemental table 3 of our previous publication.²² An ASV of *Klebsiella* accounted for nearly a third of all reads in weeks 3–4 [median 0.18% (IQR 0–65%) at 2 weeks, 24.6% (0–80%) in the 3rd, and 26.2% (0.1–66.9%) in the 4th week (*N.S.*)].

Predictors of Fecal calprotectin:

In bivariate analyses, FC was positively correlated with birth weight (Spearman's $r = 0.337$, $p < 0.001$), particularly during the 3rd ($r = 0.417$, $p = 0.009$) and 4th weeks ($r = 0.501$, $p = 0.005$), but not with gestation or post-menstrual age. There was no correlation between FC and the relative abundance of Proteobacteria or its constituent Gammaproteobacteria (or the other major phyla, Firmicutes and Actinobacteria) in stool. Interestingly, FC was positively correlated with the genus *Klebsiella* ($r = 0.207$, $p = 0.034$) but not with other Gammaproteobacteria (Table 1). FC was also correlated with the *Klebsiella* ASV described above ($r = 0.290$, $p = 0.003$).

We used mixed-effects models to identify predictors of FC across the repeated measurements. The best-fitting, parsimonious model (Table 2) showed positive associations with vaginal birth ($F = 83.29$, $p < 0.001$), Hispanic ethnicity ($F = 7.89$, $p = 0.006$), and *Klebsiella* abundance ($F = 4.54$, $p = 0.035$). The association with the dominant ASV of *Klebsiella* was even stronger ($F = 7.37$, $p = 0.008$). We also performed linear regression to identify determinants of FC at different time-points of stool collection. At 2 weeks, the best fitting model ($r^2 = 0.68$, $F = 35.76$, $p < 0.001$) showed positive effects of maternal BMI ($B = 6.08$, $t = 5.72$, $p < 0.001$) and confirmed the association with vaginal birth ($B = 157.62$, $t = 2.73$, $p = 0.01$). During the 3rd week ($r^2 = 0.8$, $F = 40.77$, $p < 0.001$), FC was associated with birth weight ($B = 0.3$, $t = 7.29$, $p < 0.001$) and Hispanic ethnicity ($B = 256.19$, $t = 3.52$, $p = 0.001$), but days of iron supplementation had a negative impact ($B = -20.8$, $t = -3.39$, $p = 0.002$). In the 4th week regression model ($r^2 = 0.61$, $F = 18.96$, $p < 0.001$), the associations

with birth weight ($B = 0.2$, $t = 4.93$, $p < 0.001$) and Hispanic ethnicity ($B = 334.69$, $t = 2.08$, $p = 0.048$) were still evident.

We also computed receiver-operating characteristics (ROCs) of *Klebsiella* abundance to identify a cut-off value that predicted high FC levels similar to those recorded in infants with mucosal inflammation ($>280 \mu\text{g/g}$ stool)¹². *Klebsiella* abundance 83% predicted FC $>280 \mu\text{g/g}$ with 72% sensitivity and 60% specificity (area under the curve, $\text{AUC} = 0.646$, $p = 0.028$; Figure 2a). In the 3rd week, *Klebsiella* abundance 76% predicted elevated FC with 89% sensitivity and 57% specificity ($\text{AUC} = 0.722$, $p = 0.047$; Figure 2b).

Predictors of the relative abundance of *Klebsiella* in stool:

Similar to the observed associations of FC, *Klebsiella* abundance was also related in mixed-effects models to vaginal birth ($F = 5.68$, $p = 0.015$) and Hispanic ethnicity ($F = 124.33$, $p < 0.001$). The dominant ASV of *Klebsiella* was also affected by the same clinical antecedents, vaginal birth ($F = 13.46$, $p < 0.001$) and Hispanic ethnicity ($F = 57.81$, $p < 0.001$).

Clustering of infants based on *Klebsiella* abundance at 2 weeks:

To investigate whether the variability (range 0–98%) in *Klebsiella* abundance we observed in the 2-week stool samples indicated the presence of more than one distinct subgroups of infants in our cohort, we used agglomerative hierarchical and K-means clustering procedures. These analyses revealed 2 subgroups with distinct cluster centers (13% vs. 94% in clusters 1 and 2, respectively; $F = 370.83$, $p < 0.001$; analysis of variance; Figure 2). In cluster 1 ($n = 23$), the relative abundance of *Klebsiella* (and its dominant ASV) started at low levels and increased over time (Friedman test, $p = 0.009$ and 0.021 for *Klebsiella* and its dominant ASV, respectively). In contrast, *Klebsiella* dominated the intestinal microbiome in cluster 2 ($n = 21$). There was minimal loss of *Klebsiella* abundance over time ($p = 0.018$; Table 3) but the *Klebsiella* ASV remained stable. Gneiss (balance tree analysis)²⁰ confirmed that these cluster differences in *Klebsiella* abundance were significant and not merely due to decreased abundance of other taxa (Figure 3). In support of these data, the accumulative ratios of major bacterial families vs. *Klebsiella* (Figure 4) decreased over time in cluster 1, consistent with the acquisition of *Klebsiella*. Cluster 2 started with and maintained low ratios, indicating persistently high levels of *Klebsiella* compared to other taxa.

The two clusters were similar in birth weight and gestational age. Cluster 1 experienced a shorter duration of maternal ruptured membranes but had more days on antibiotics. There was no difference in major neonatal morbidities, days to full enteral feeds (150 mL/kg), days on positive pressure respiratory support, Z-scores for weight on postnatal day 28 and at 36 weeks' post-menstrual age (based on Fenton growth charts), and length of hospital stay. A summary of the clinical characteristics of the two clusters in Table 4.

In mixed-effects models, *Klebsiella* abundance in cluster 1 increased with postnatal age ($F = 5.0$, $p = 0.028$) and antenatal steroids ($F = 5.8$, $p = 0.019$), whereas antenatal magnesium sulfate ($F = 8.5$, $p = 0.005$) and chorioamnionitis ($F = 5.2$, $p = 0.026$) had a negative effect. The association with antenatal steroids was also detectable in cluster 2 ($F = 5.3$, $p = 0.025$). In cluster 1, *Klebsiella* ASV also increased with postnatal age ($F = 5.8$, $p = 0.018$), and decreased with antenatal magnesium sulfate ($F = 8.1$, $p = 0.006$) and chorioamnionitis ($F =$

4.6, $p = 0.035$). In cluster 2, *Klebsiella* ASV was associated with vaginal delivery ($F = 6.6$, $p = 0.013$).

Cluster 1 had lower FC levels overall than cluster 2 [median (IQR) 148 (99–215) in cluster 1 vs. 226 (112–362) $\mu\text{g/g}$ in cluster 2, respectively; $p = 0.011$]. FC values were the highest in cluster 2 at 2 weeks [median (IQR) 243 (121–331) $\mu\text{g/g}$ in cluster 2 vs. 148 (83–199) in cluster 1; $p = 0.04$]. Antenatal magnesium sulfate had a negative effect on FC in cluster 1 ($F = 15.0$, $p < 0.001$), but a positive effect was seen in cluster 2 ($F = 5.8$, $p = 0.02$). Antenatal steroids had a negative effect on FC in cluster 2 ($F = 6.0$, $p = 0.018$).

DISCUSSION

We present a detailed investigation of the relationship between fecal colonization with Gammaproteobacteria in VLBW infants and FC, a biomarker of gut mucosal inflammation. In our cohort, Gammaproteobacteria abundance did not affect FC expression. Instead, we found FC to be associated specifically with the presence of *Klebsiella*, and even more strongly, with a particular ASV within this genus. We observed two distinct paths to enteric colonization with *Klebsiella*: infants in cluster 1 began with very low abundance and gained modest levels of these bacteria over time, whereas the gut microbiome in cluster 2 was completely dominated by *Klebsiella*. High FC in cluster 2 infants indicated ongoing host-microbial interaction and a measurable innate immune response to these luminal bacteria. Importantly, *Klebsiella* abundance 83% predicted elevated FC levels similar to those seen during intestinal inflammation and NEC. To our knowledge, this is the first study to evaluate enteric dysbiosis in preterm infants in the context of the host inflammatory response.

The association of FC with fecal *Klebsiella* in our study raises several scientific questions. First, our findings call into question the practice of evaluating Gammaproteobacteria *en bloc* in the preterm microbiome. In recent years, several studies show that premature infants with a diagnosis of NEC had developed enteric dysbiosis with increased Gammaproteobacteria for several days to weeks antedating NEC onset (5–8). However, the lack of correlation between Gammaproteobacteria abundance and FC in our cohort suggests that Gammaproteobacteria may be too diverse a group to consistently exert a net inflammatory effect, and indicates a need for defining dysbiosis at higher levels of resolution. Second, we recorded highest FC levels in cluster 2 during the early neonatal period, when *Klebsiella* dominated in a setting of low alpha-diversity. These findings are important in efforts to develop a clinically-relevant definition for dysbiosis; further study is needed to clarify whether the relative abundance of a pathobiont is relevant in isolation, or if it needs to be interpreted in conjunction with the diversity indices. Third, further study is needed to determine whether the early appearance of *Klebsiella* in some premature infants was unique to our center or if it represents a universal phenomenon in 13 of 36 these patients. The inflammatory effects of *Klebsiella* in the intestine are plausible, considering the presence of potent virulence factors such as cell wall components and enterotoxins (23,24). *Klebsiella* are recognized intestinal pathogens of preterm and term neonates, having been identified in diarrhea, ecchymotic colitis, bacteremia during NEC and short-bowel syndrome, and even in NEC outbreaks (25–29). The correlation between fecal *Klebsiella* and elevated FC has been previously noted in infantile colic (30). Early colonization with *Klebsiella* has been noted in

other preterm cohorts (6,8), but the possibility of finding distinct inflammation-driving pathobiont(s) at other centers cannot be excluded. In our own cohort, the correlation between FC and *Klebsiella* was relatively modest ($r = 0.21$), and when clustering was attempted based on FC, subgroup differences in *Klebsiella* did not reach significance. Finally, even though current evidence supports dysbiosis as an inflammatory stimulus, we need to acknowledge that a reciprocal effect, where pre-existing inflammation in the intestinal mucosa may promote a Gammaproteobacterial bloom, also warrants exploration.

Calprotectin is a 24 kDa dimer of the calcium-binding proteins calgranulin A/S100A8 and calgranulin B/S100A9, expressed in neutrophils and inflammatory monocytes (11–14). FC has found wide acceptance as a non-invasive marker of gut mucosal inflammation because it is strongly correlated with leukocyte infiltration in the mucosa, is stable in stool at room temperature for several days, and is easy to measure (31). Premature infants tend to have higher and more variable FC than adults, but higher cut-off values such as 280 $\mu\text{g/g}$ stool have been correlated with intestinal inflammation and NEC (11–14). Treatment with non-steroidal anti-inflammatory drugs (NSAIDs) can increase FC ³², but in our cohort, the FC levels in the 2 infants who received indomethacin for treatment of PDA were not higher than the rest.

Vaginal birth was an important determinant of the relative abundance of *Klebsiella* (and the dominant *Klebsiella*) in stool and of FC expression. These findings suggest that *Klebsiella* were likely acquired via vertical, mother-to-infant transmission. The near-total dominance of the gut microbiome in some infants (cluster 2) within the first 2 weeks after birth also lends credence to this possibility. However, the source of these bacteria is unclear. During pregnancy, the vaginal microbiome is dominated by Firmicutes (33,34). In women with vaginal dysbiosis, pathobionts such as *Prevotella*, *Sneathia*, *Atopobium*, *Mycoplasma*, and *Gardnerella* can be identified, but *Klebsiella* are infrequent (34). Perinatal exposure to maternal enteric flora during vaginal birth and to the microbiome of human milk, which contain *Klebsiella* (35,36), are two possible sources. Increased maternal BMI and diabetes have been previously associated with enrichment of *Enterobacteriaceae*, including *Klebsiella*, in both maternal and neonatal gut microbiome (37). Unfortunately, we are unable to investigate these findings further because we did not collect maternal samples in the present study.

We found higher *Klebsiella* abundance and FC in Hispanic infants. Although racial/ethnic differences in the enteric microbiome are possible (38), the type of feedings could have had a confounding influence: 8/9 (88.8%) Hispanic infants received only human milk vs. 17/36 (47.2%) infants of other ethnic groups ($p = 0.03$). The correlation between birth weight and FC likely reflects the maturational expansion of the mucosal leukocyte populations, and thus, the improving ability of the host to launch an innate immune response. Associations with antenatal steroids and magnesium sulfate may be more difficult to explain; steroids could influence the host-microbial cross-talk through altered leukocyte function and cytokine expression, and epigenetic changes in the mucosa (39). Perinatal exposure to magnesium sulfate can also alter gut motility and some immune responses (40), but the effects on fecal microbiome are unclear.

In conclusion, we show that FC expression in VLBW infants was specifically associated with the genus *Klebsiella* and a dominant ASV mapping to this genus, but not with the relative abundance of all Gammaproteobacteria in stool. These findings are important in the quest to develop a clinically-relevant definition of enteric dysbiosis in premature infants. Major strengths of our study are repeated measurements of the gut microbiome and FC, and the single-patient room floor-plan of our NICU, which now represents the layout in most of the newer NICUs in the United States. There are important limitations, such as the limited sample size and the single-center design. In our cohort, *Klebsiella* abundance affected FC but not the clinical outcome; we had only a single patient with NEC, which made it difficult to analyze the clinical effects of abnormal microbial colonization and mucosal inflammation. Further study is needed in a larger, multicenter cohort to validate our findings, particularly to confirm the dichotomous development of the gut microbiome and the importance of *Klebsiella* as a leading pathobiont in preterm infants.

Acknowledgements:

This research was funded by National Institutes of Health awards NR015446 (MG), HL124078 (AM) and HL133022 (AM) and T32GM007281 (ALY).

The authors thank the research nurses, Judy Zaritt and Marcia Kneusel, for identification of eligible infants, and collection of clinical information and stool samples. We are also indebted to the nursing staff at the NICU at TGH for their constant support.

Statement of Financial Supports: This research was funded by National Institutes of Health awards NR015446 (MG), HL124078 (AM) and HL133022 (AM) and T32GM007281 (ALY).

REFERENCES

1. Penders J, Thijs C, Vink C, et al. Factors influencing the composition of the intestinal microbiota in early infancy. *Pediatrics* 2006;118:511–21. [PubMed: 16882802]
2. Turnbaugh PJ, Hamady M, Yatsunenko T, et al. A core gut microbiome in obese and lean twins. *Nature* 2009;457:480–4. [PubMed: 19043404]
3. Sommer F, Backhed F. The gut microbiota--masters of host development and physiology. *Nat Rev Microbiol* 2013;11:227–38. [PubMed: 23435359]
4. Stewart CJ, Embleton ND, Marrs ECL, et al. Longitudinal development of the gut microbiome and metabolome in preterm neonates with late onset sepsis and healthy controls. *Microbiome* 2017;5:75. [PubMed: 28701177]
5. Wang Y, Hoenig JD, Malin KJ, et al. 16S rRNA gene-based analysis of fecal microbiota from preterm infants with and without necrotizing enterocolitis. *ISME J* 2009;3:944–54. [PubMed: 19369970]
6. Warner BB, Deych E, Zhou Y, et al. Gut bacteria dysbiosis and necrotizing enterocolitis in very low birthweight infants: a prospective case-control study. *Lancet* 2016.
7. Pammi M, Cope J, Tarr PI, et al. Intestinal dysbiosis in preterm infants preceding necrotizing enterocolitis: a systematic review and meta-analysis. *Microbiome* 2017;5:31. [PubMed: 28274256]
8. Torrazza RM, Ukhanova M, Wang X, et al. Intestinal microbial ecology and environmental factors affecting necrotizing enterocolitis. *PLoS One* 2013;8:e83304. [PubMed: 24386174]
9. Goyal MS, Venkatesh S, Milbrandt J, Gordon JI, Raichle ME. Feeding the brain and nurturing the mind: Linking nutrition and the gut microbiota to brain development. *Proc Natl Acad Sci U S A* 2015;112:14105–12. [PubMed: 26578751]
10. Smith PD, Smythies LE, Mosteller-Barnum M, et al. Intestinal macrophages lack CD14 and CD89 and consequently are down-regulated for LPS- and IgA-mediated activities. *J Immunol* 2001;167:2651–6. [PubMed: 11509607]

11. MacQueen BC, Christensen RD, Yost CC, et al. Elevated fecal calprotectin levels during necrotizing enterocolitis are associated with activated neutrophils extruding neutrophil extracellular traps. *J Perinatol* 2016;36:862–9. [PubMed: 27388941]
12. Zhang M, Zhang X, Zhang J. Diagnostic Value of Fecal Calprotectin in Preterm Infants with Necrotizing Enterocolitis. *Clin Lab* 2016;62:863–9. [PubMed: 27349012]
13. Rouge C, Butel MJ, Piloquet H, et al. Fecal calprotectin excretion in preterm infants during the neonatal period. *PLoS One* 2010;5:e11083. [PubMed: 20552029]
14. Albanna EA, Ahmed HS, Awad HA. Stool calprotectin in necrotizing enterocolitis. *J Clin Neonatol* 2014;3:16–9. [PubMed: 24741535]
15. Moore HM, Kelly A, McShane LM, Vaught J. Biospecimen Reporting for improved study quality (BRISQ). *Transfusion* 2013;53:e1. [PubMed: 23844646]
16. Walters W, Hyde ER, Berg-Lyons D, et al. Improved Bacterial 16S rRNA Gene (V4 and V4–5) and Fungal Internal Transcribed Spacer Marker Gene Primers for Microbial Community Surveys. *mSystems* 2016;1.
17. Torrazza RM, Li N, Neu J. Decoding the enigma of necrotizing enterocolitis in premature infants. *Pathophysiology* 2014;21:21–7. [PubMed: 24373685]
18. Fullerton BS, Sparks EA, Morrow KA, et al. Hospital transfers and patterns of mortality in very low birth weight neonates with surgical necrotizing enterocolitis. *J Pediatr Surg* 2016;51:932–5. [PubMed: 27230800]
19. Zhou Y, Shan G, Sodergren E, Weinstock G, Walker WA, Gregory KE. Longitudinal analysis of the premature infant intestinal microbiome prior to necrotizing enterocolitis: a case-control study. *PLoS One* 2015;10:e0118632. [PubMed: 25741698]
20. Morton JT, Sanders J, Quinn RA, et al. Balance Trees Reveal Microbial Niche Differentiation. *mSystems* 2017;2.
21. Ho TBT, Groer MW, Kane B, et al. Dichotomous Development of the Gut Microbiome in Preterm Infants. *Microbiome* 2018;in review.
22. Ho TTB, Groer MW, Kane B, et al. Dichotomous development of the gut microbiome in preterm infants. *Microbiome* 2018;6:157. [PubMed: 30208950]
23. Lu MC, Chen YT, Chiang MK, et al. Colibactin Contributes to the Hypervirulence of pks(+) K1 CC23 *Klebsiella pneumoniae* in Mouse Meningitis Infections. *Front Cell Infect Microbiol* 2017;7:103. [PubMed: 28409125]
24. Straus DC, Lonon MK, Woods DE, Garner CW. Production of an extracellular toxic complex by various strains of *Pseudomonas cepacia*. *J Med Microbiol* 1989;30:17–22. [PubMed: 2778792]
25. Gregersen N, Van Nierop W, Von Gottberg A, Duse A, Davies V, Cooper P. *Klebsiella pneumoniae* with extended spectrum beta-lactamase activity associated with a necrotizing enterocolitis outbreak. *Pediatr Infect Dis J* 1999;18:963–7. [PubMed: 10571430]
26. Hill HR, Hunt CE, Matsen JM. Nosocomial colonization with *Klebsiella*, type 26, in a neonatal intensive-care unit associated with an outbreak of sepsis, meningitis, and necrotizing enterocolitis. *J Pediatr* 1974;85:415–9. [PubMed: 4610423]
27. Boccia D, Stolfi I, Lana S, Moro ML. Nosocomial necrotising enterocolitis outbreaks: epidemiology and control measures. *Eur J Pediatr* 2001;160:385–91. [PubMed: 11421422]
28. Stone HH, Kolb LD, Geheber CE. Bacteriologic considerations in perforated necrotizing enterocolitis. *South Med J* 1979;72:1540–4. [PubMed: 390716]
29. Canioni D, Pauliat S, Gaillard JL, et al. Histopathology and microbiology of isolated rectal bleeding in neonates: the so-called ‘ecchymotic colitis’. *Histopathology* 1997;30:472–7. [PubMed: 9181369]
30. Rhoads JM, Fatheree NY, Norori J, et al. Altered fecal microflora and increased fecal calprotectin in infants with colic. *J Pediatr* 2009;155:823–8 e1. [PubMed: 19628216]
31. Roseth AG, Schmidt PN, Fagerhol MK. Correlation between faecal excretion of indium-111-labelled granulocytes and calprotectin, a granulocyte marker protein, in patients with inflammatory bowel disease. *Scand J Gastroenterol* 1999;34:50–4. [PubMed: 10048733]
32. Gisbert JP, McNicholl AG. Questions and answers on the role of faecal calprotectin as a biological marker in inflammatory bowel disease. *Dig Liver Dis* 2009;41:56–66. [PubMed: 18602356]

33. Stout MJ, Zhou Y, Wylie KM, Tarr PI, Macones GA, Tuuli MG. Early pregnancy vaginal microbiome trends and preterm birth. *American journal of obstetrics and gynecology* 2017.
34. Kindinger LM, Bennett PR, Lee YS, et al. The interaction between vaginal microbiota, cervical length, and vaginal progesterone treatment for preterm birth risk. *Microbiome* 2017;5:6. [PubMed: 28103952]
35. Urbaniak C, Angelini M, Gloor GB, Reid G. Human milk microbiota profiles in relation to birthing method, gestation and infant gender. *Microbiome* 2016;4:1. [PubMed: 26739322]
36. Patel SH, Vaidya YH, Patel RJ, Pandit RJ, Joshi CG, Kunjadiya AP. Culture independent assessment of human milk microbial community in lactational mastitis. *Scientific reports* 2017;7:7804. [PubMed: 28798374]
37. Verdam FJ, Fuentes S, de Jonge C, et al. Human intestinal microbiota composition is associated with local and systemic inflammation in obesity. *Obesity (Silver Spring)* 2013;21:E607–15. [PubMed: 23526699]
38. Gupta VK, Paul S, Dutta C. Geography, Ethnicity or Subsistence-Specific Variations in Human Microbiome Composition and Diversity. *Front Microbiol* 2017;8:1162. [PubMed: 28690602]
39. Cortese R, Lu L, Yu Y, Ruden D, Claud EC. Epigenome-Microbiome crosstalk: A potential new paradigm influencing neonatal susceptibility to disease. *Epigenetics* 2016;11:205–15. [PubMed: 26909656]
40. Ghidini A, Espada RA, Spong CY. Does exposure to magnesium sulfate in utero decrease the risk of necrotizing enterocolitis in premature infants? *Acta Obstet Gynecol Scand* 2001;80:126–9. [PubMed: 11167206]

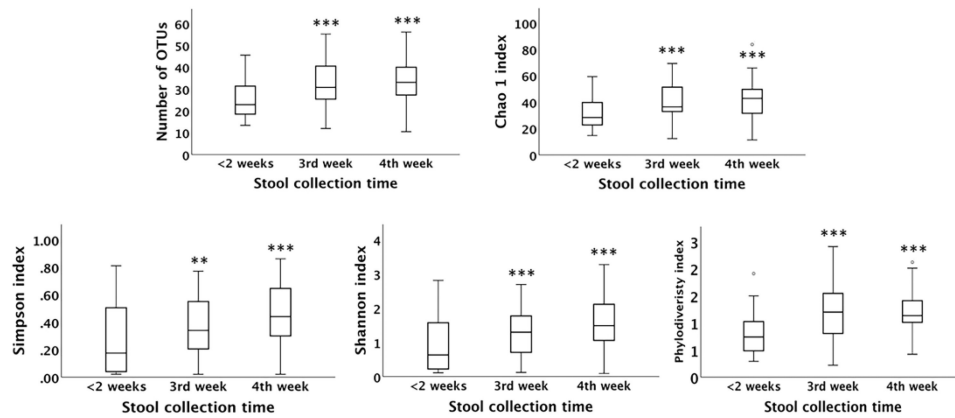


Figure 1: Longitudinal change in alpha-diversity indices.

Boxplots (median and 95% confidence interval) summarize number of OTUs, Chao 1, Simpson, Shannon, and phylogenetic indices. (**) and (***) indicate $p < 0.01$ and $p < 0.001$ respectively in comparison to the < 2 weeks samples.

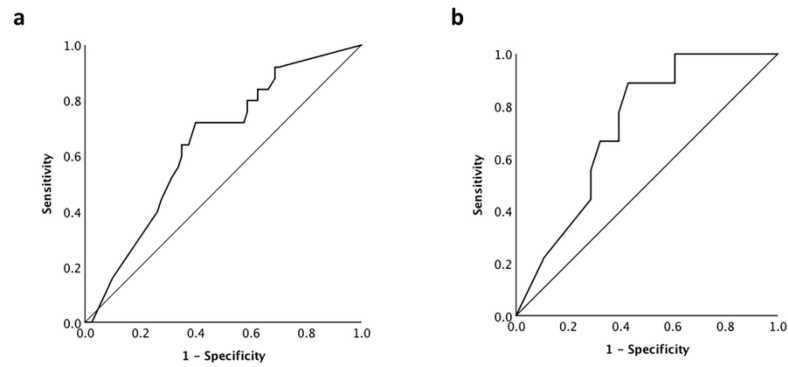


Figure 2: Fecal *Klebsiella* predict FC expression:

(a) Relative abundance of *Klebsiella* 83% in stool predicted high FC (>280 $\mu\text{g/g}$ of stool) with 72% sensitivity and 60% specificity (area under curve, AUC = 0.646); (b) During the 3rd postnatal week, *Klebsiella* abundance 76% predicted high FC with 89% sensitivity and 57% specificity (AUC = 0.722).

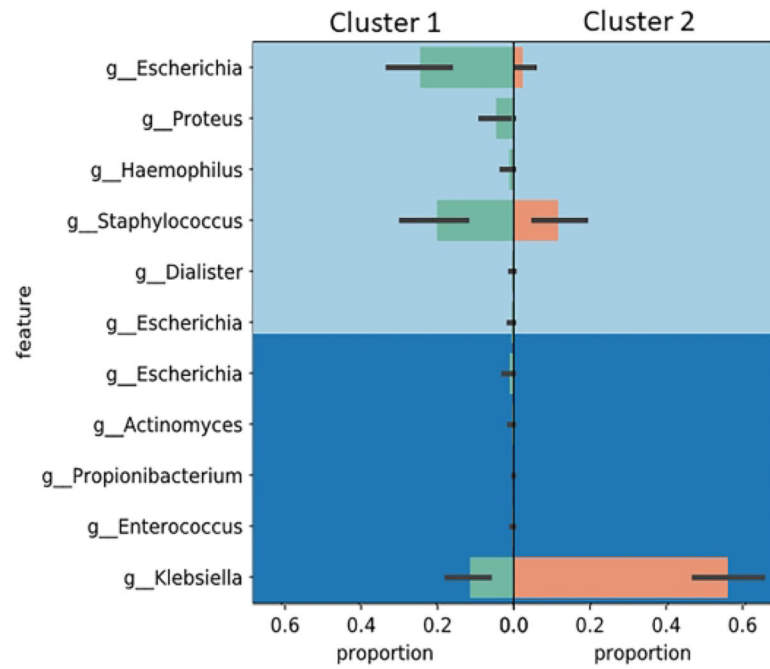


Figure 3:
Proportion plot based on balance tree calculation shows shifts in major genera between the two clusters.

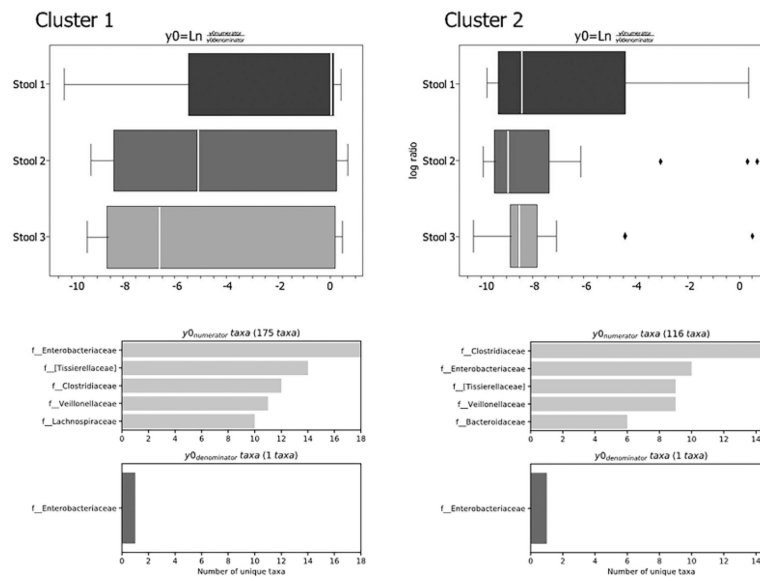


Figure 4: Temporal change in bacterial abundances at family level relative to *Klebsiella* in the two clusters.

Top: Boxplots summarize accumulative ratios of major bacterial families *vs. Klebsiella* (natural log-transformed, Ln) in the two patient clusters. *Bottom:* Bar diagrams summarize the distribution of major bacterial taxa at the family level in the numerator and the denominator.

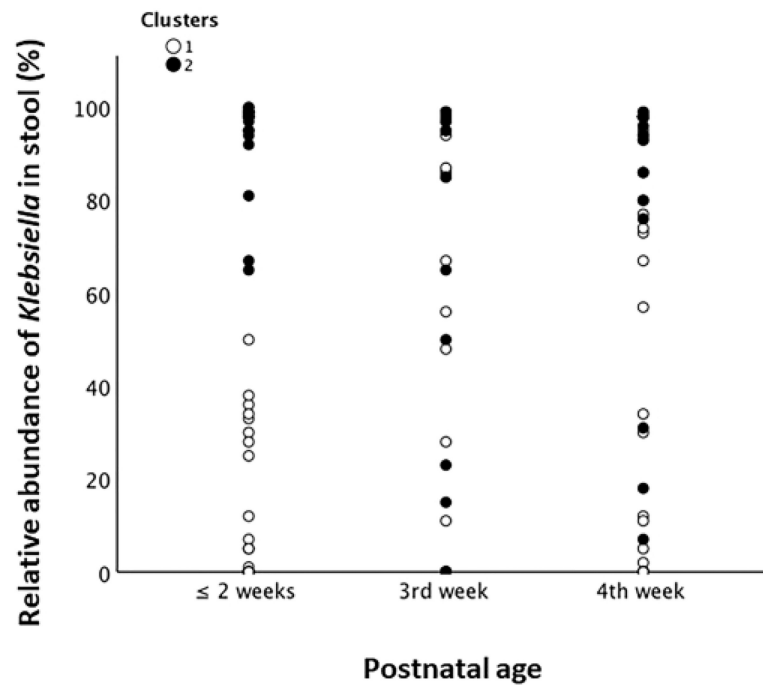


Figure 5: Longitudinal change in fecal *Klebsiella* abundance.

Scatter dot plots summarize fecal *Klebsiella* abundance over time. Open and solid circles indicate the two sub-groups of infants identified in our cohort by K-means clustering analysis of fecal *Klebsiella* abundance at 2 weeks. Cluster 1 (23 infants, open circles) started with low *Klebsiella* abundance and gained these bacteria over time (Friedman test, $p = 0.009$). Cluster 2 (21 infants, solid circles) showed high *Klebsiella* abundance with minor attrition in weeks 3 and 4 ($p = 0.018$).

Table 1.

Bivariate correlation between FC and the relative abundance of major genera in the class Gammaproteobacteria

Genus	r	p-value
<i>Klebsiella</i>	0.207	0.034
<i>Haemophilus</i>	−0.023	0.820
<i>Proteus</i>	−0.099	0.316
<i>Acinetobacter</i>	0.089	0.365
<i>Vibrio</i>	−0.142	0.148
<i>Pseudomonas</i>	0.002	0.985
<i>Enhydrobacter</i>	−0.040	0.685
<i>Salmonella</i>	−0.084	0.392
<i>Microbulbifer</i>	−0.103	0.297
<i>Trabulsiella</i>	0.176	0.073
<i>Arenimonas</i>	0.024	0.804
<i>Photobacterium</i>	−0.041	0.680
<i>Cardiobacterium</i>	−0.041	0.680
<i>Serratia</i>	0.019	0.845
<i>Stenotrophomonas</i>	−0.064	0.514
<i>Morganella</i>	−0.007	0.946
<i>Erwinia</i>	−0.057	0.566

Table 2.

Predictors of FC

Parameter	Estimate	Std. Error	df	t	p-value	95% Confidence Interval	
						Lower Bound	Upper Bound
Vaginal birth	348.220	67.270	105.000	5.176	0.000	214.836	481.604
Non-Hispanic ethnicity	-127.186	51.705	105.000	-2.460	0.016	-229.707	-24.664
Relative abundance of Klebsiella	0.932	0.437	105.000	2.132	0.035	0.065	1.799
Female gender	-61.289	36.944	105.000	-1.659	0.100	-134.542	11.964

Table 3:

Klebsiella abundances in stool and FC over time, by cluster

Time of stool collection	Cluster 1			Cluster 2		
	<i>Klebsiella</i> (%) [median (IQR)]	<i>Klebsiella</i> ASV (%) [median (IQR)]	FC (µg/g stool) [median(IQR)]	<i>Klebsiella</i> (%) [median (IQR)]	<i>Klebsiella</i> ASV (%) [median (IQR)]	FC (µg/g stool) [median (IQR)]
2 weeks	5 (0–30)	0 (0–0.1)	148 (83–199)	98 (94.5–99)	66.4 (0.7–94)	243 (121–331)
3 rd week	19.5 (0–86)	0.1 (0–36.6)	162 (123–210)	98 (70–98.8)	80.7 (15.3–89.4)	227 (142–417)
4 th week	21 (0–74.8)	0.3 (0–26.2)	134 (79–285)	96 (80–98)	59.6 (29.4–75.4)	163 (102–436)

Table 4:

Perinatal and Neonatal Clinical Characteristics of the Two Clusters

Characteristic	Cluster 1 N=23	Cluster 2 N=21	p values
Gestational age (weeks, mean \pm SD)	28 \pm 2	28 \pm 2	0.281
Birth weight, grams, (mean, SD)	1099 \pm 220	1144 \pm 187	0.211
Male	12 (52.2%)	8 (38.1%)	0.382
Hispanic ethnicity	4 (17.4%)	5 (23.8%)	0.716
Race			0.302
Black	12 (52.2%)	7 (33.3%)	
White	11 (47.8%)	13 (61.9%)	
Antenatal medications			
Steroids	20 (87%)	19 (90.5%)	1.000
Magnesium	17 (73.9%)	17 (81%)	0.724
Vaginal birth	4 (17.4%)	7 (33.3%)	0.303
Multiple birth	6 (26.1%)	1 (4.8%)	0.097
Ruptured membrane (hours, mean \pm SD)	17\pm64	65\pm173	0.038
Chorioamnionitis	15 (65.2%)	10 (47.6%)	0.521
Maternal hypertension	6 (26.1%)	7 (33.3%)	0.744
Small for gestational age	2 (8.7%)	1 (4.8%)	1.000
Respiratory distress syndrome	14 (60.9%)	26 (57.8%)	0.762
Oxygen on day 28	6 (26.1%)	4 (19.0%)	0.724
Oxygen on day 36	1 (8.7%)	1 (4.8%)	1.000
Significant patent ductus arteriosus	2 (13.3%)	3 (14.3%)	0.658
Indomethacin	0	2 (9.5%)	0.222
Patent ductus arteriosus ligation	1 (2.2%)	0	1.000
Retinopathy of prematurity	2 (8.7%)	4 (19%)	0.403
Intraventricular hemorrhage	1 (4.3%)	4 (19%)	0.176
Necrotizing enterocolitis	1 (4.3%)	0	1.000
Days on antibiotics, (mean, SD)	5.3 (6.0)	3.7 (2.5)	0.045
Positive blood culture	3 (13.0%)	2 (9.5%)	1.000

Characteristic	Cluster 1 N=23	Cluster 2 N=21	p values
Packed red blood cell transfusion	7 (30.4%)	8 (38.1%)	0.752
Feeding type			0.936
Maternal breast milk only	12 (52.2%)	12 (57.1%)	
Formula only	1 (4.3%)	1 (4.8%)	
Mixed feeding types	10 (43.5%)	8 (38.1%)	
Days to full enteral feeds; median (range)	11 (6–19)	11 (5–17)	0.623
Z-score for weight, postnatal day 28; median (range)	−0.89 (−3.72 to 1.14)	−0.52 (−2.56 to 1.71)	0.301
Z-score for weight, post-menstrual age 36 weeks; median (range)	−0.81 (−3.72 to 1.09)	−0.41 (−2.39 to 1.56)	0.459
Days on respiratory support	34 (0–92)	42 (0–68)	0.689
Discharge weight <10 th percentile	5 (21.7%)	6 (28.6%)	0.732
Length of stay (days, mean ± SD)	71±34	64±21	0.163